

RECEIVED
FEB 25 2002
TECH CENTER 1600/2900

1645



P#19

1600

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/596,101C

DATE: 02/12/2002
TIME: 11:43:16

Input Set : A:\EP.txt
Output Set: N:\CRF3\02122002\I596101C.raw

3 <110> APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL
4 de Baetselier, Patrick
5 Beschin, Alain
7 <120> TITLE OF INVENTION: Peptides and nucleic acids derived from Eisenia foetida and
the use
8 thereof

10 <130> FILE REFERENCE: 2676-4432US
12 <140> CURRENT APPLICATION NUMBER: US 09/596,101C
13 <141> CURRENT FILING DATE: 2000-06-16
15 <150> PRIOR APPLICATION NUMBER: PCT/EP98/08169
16 <151> PRIOR FILING DATE: 1998-12-16
18 <150> PRIOR APPLICATION NUMBER: EP 97203974.7
19 <151> PRIOR FILING DATE: 1997-12-17
21 <160> NUMBER OF SEQ ID NOS: 15
23 <170> SOFTWARE: PatentIn version 3.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 13
27 <212> TYPE: PRT
28 <213> ORGANISM: Eisenia foetida

W--> 29 <400> SEQUENCE: 1

31 Ser Gly Glu Ile Asp Ile Ile Glu Thr Ile Gly Asn Arg
32 1 5 10

35 <210> SEQ ID NO: 2
36 <211> LENGTH: 1155
37 <212> TYPE: DNA

38 <213> ORGANISM: Eisenia foetida

40 <220> FEATURE:

41 <221> NAME/KEY: CDS
42 <222> LOCATION: (1)..(1152)
43 <223> OTHER INFORMATION:
45 <220> FEATURE:

46 <221> NAME/KEY: sig_peptide
47 <222> LOCATION: (1)..(51)
48 <223> OTHER INFORMATION:

50 <220> FEATURE:

51 <221> NAME/KEY: mat_peptide
52 <222> LOCATION: (52)..()
53 <223> OTHER INFORMATION:

55 <400> SEQUENCE: 2

56 atg agg tgg act ttg gtg gtc ttg tgt ctg ctg ttc ggt gaa ggg ttc
57 Met Arg Trp Thr Leu Val Val Cys,Leu Leu Phe Gly Glu Gly Phe
58 -15 -10 -5
60 gcc ttc acc gac tgg gat caa tat cac atc gtc tgg cag gac gaa ttc
61 Ala Phe Thr Asp Trp Asp Gln Tyr His Ile Val Trp Gln Asp Glu Phe

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/596,101C

DATE: 02/12/2002
TIME: 11:43:16

Input Set : A:\EP.txt
Output Set: N:\CRF3\02122002\I596101C.raw

62	-1	1	5	10	15												
64	gat	tac	ttt	gat	ggc	gcg	aag	tgg	caa	cat	gag	gtc	aca	gca	act	ggc	144
65	Asp	Tyr	Phe	Asp	Gly	Ala	Lys	Trp	Gln	His	Glu	Val	Thr	Ala	Thr	Gly	
66																30	
68	gga	ggg	aac	agc	gaa	ttc	caa	ctg	tac	aca	cag	gat	ggg	gcc	aac	agc	192
69	Gly	Gly	Asn	Ser	Glu	Phe	Gln	Leu	Tyr	Thr	Gln	Asp	Gly	Ala	Asn	Ser	
70																	
72	ttc	gtt	cga	gat	gga	aag	ttt	ttc	att	aag	ccg	acg	ttg	ctg	gct	gac	240
73	Phe	Val	Arg	Asp	Gly	Lys	Leu	Phe	Ile	Lys	Pro	Thr	Leu	Leu	Ala	Asp	
74																	
76	aac	atc	aac	cca	cag	acg	ggt	gcg	cca	ttt	gga	acc	gat	ttc	atg	tac	288
77	Asn	Ile	Asn	Pro	Gln	Thr	Gly	Ala	Pro	Phe	Gly	Thr	Asp	Phe	Met	Tyr	
78																	
80	aat	gga	gtt	cta	gat	gtc	tgg	gct	atg	tac	ggg	gcc	tgc	acg	aat	acg	336
81	Asn	Gly	Val	Leu	Asp	Val	Trp	Ala	Met	Tyr	Gly	Ala	Cys	Thr	Asn	Thr	
82	80															95	
84	gac	aac	aac	gga	tgc	tac	agg	acg	gga	gcc	gct	gcc	gac	att	cca	ccg	384
85	Asp	Asn	Asn	Gly	Cys	Tyr	Arg	Thr	Gly	Ala	Ala	Gly	Asp	Ile	Pro	Pro	
86																	
88	100															105	110
89	gcc	atg	tcg	gca	cga	gtt	cga	acc	ttc	cag	aaa	tac	acg	ttc	acc	cac	432
90	Ala	Met	Ser	Ala	Arg	Val	Arg	Thr	Phe	Gln	Lys	Tyr	Ser	Phe	Thr	His	
92																	
93	gga	cgc	gtt	gtc	gtt	cac	gcc	aag	atg	ccc	gtc	gga	gac	tgg	ctc	tgg	480
94	Gly	Arg	Val	Val	Val	His	Ala	Lys	Met	Pro	Val	Gly	Asp	Trp	Leu	Trp	
96																	
97	cca	gcc	att	tgg	atg	ttg	ccg	gag	gat	tgg	gtc	tat	ggc	gga	tgg	cct	528
98	Pro	Ala	Ile	Trp	Met	Leu	Pro	Glu	Asp	Trp	Val	Tyr	Gly	Gly	Trp	Pro	
100																	
101	cga	tcg	ggc	gag	atc	gac	atc	att	gaa	aca	atc	ggc	aac	cga	gat	ttc	576
102	Arg	Ser	Gly	Glu	Ile	Asp	Ile	Ile	Glu	Thr	Ile	Gly	Asn	Arg	Asp	Phe	
104																	
105	aag	aac	act	ggt	gga	gag	ttc	ctt	gga	att	cag	aag	atg	gga	tca	acg	624
106	Lys	Asn	Thr	Gly	Gly	Glu	Phe	Leu	Gly	Ile	Gln	Lys	Met	Gly	Ser	Thr	
108																	
109	atg	cac	tgg	ggt	cca	gga	tgg	gac	gac	aac	cga	tac	tgg	ctg	acc	acg	672
110	Met	His	Trp	Gly	Pro	Gly	Trp	Asp	Asp	Asn	Arg	Tyr	Trp	Leu	Thr	Ser	
112																	
113	ctt	ccg	aaa	cac	tca	gac	gat	tgg	aac	tac	ggt	gac	aac	ttc	cac	acg	720
114	Leu	Pro	Lys	His	Ser	Asp	Asp	Trp	Asn	Tyr	Gly	Asp	Asn	Phe	His	Thr	
116																	
117	ttc	tgg	tgc	gac	tgg	agt	ccc	aac	gga	ctg	agg	ttc	ttc	gta	gac	gac	768
118	Phe	Trp	Phe	Asp	Trp	Ser	Pro	Asn	Gly	Leu	Arg	Phe	Phe	Val	Asp	Asp	
120																	
121	gag	aac	cag	gct	ctg	ctc	gat	gtt	cct	tat	cct	ctg	att	gat	gcg	aat	816
122	Glu	Asn	Gln	Ala	Leu	Leu	Asp	Val	Pro	Tyr	Pro	Leu	Ile	Asp	Ala	Asn	
124																	
125	240															255	
126	cca	tgg	tgg	gtg	gat	ttc	tgg	gag	tgg	gga	aag	ccc	tgg	ctt	cct	caa	864
126	Pro	Trp	Trp	Val	Asp	Phe	Trp	Glu	Trp	Gly	Lys	Pro	Trp	Leu	Pro	Gln	
126																	
126	260																
126	265																
126	270																

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/596,101C

DATE: 02/12/2002
TIME: 11:43:16

Input Set : A:\EP.txt
Output Set: N:\CRF3\02122002\I596101C.raw

128 tac gaa aat gac aat cca tgg gct gga gga acg aac ctg gct ccc ttc	912
129 Tyr Glu Asn Asp Asn Pro Trp Ala Gly Gly Thr Asn Leu Ala Pro Phe	
130 275 280 285	
132 gac caa aat ttc cac ttc att ctg aac gtg gct gtc gga gga acg aac	960
133 Asp Gln Asn Phe His Phe Ile Leu Asn Val Ala Val Gly Gly Thr Asn	
134 290 295 300	
136 ggc ttc atc ccg gac ggt tgc atc aat cgc ggc gga gac ccg gcc ctg	1008
137 Gly Phe Ile Pro Asp Gly Cys Ile Asn Arg Gly Gly Asp Pro Ala Leu	
138 305 310 315	
140 cag aag ccg tgg agc aat ggg gac tgg tac aac gat gca atg agg aaa	1056
141 Gln Lys Pro Trp Ser Asn Gly Asp Trp Tyr Asn Asp Ala Met Arg Lys	
142 320 325 330 335	
144 ttc ttc gac gcc aga gga aac tgg aag tgg acg tgg gat gac gag gga	1104
145 Phe Phe Asp Ala Arg Gly Asn Trp Lys Trp Thr Trp Asp Asp Glu Gly	
146 340 345 350	
148 gac aac aat gcc atg cag gtc gat tac atc cga gtc tac aag cgc aac	1152
149 Asp Asn Asn Ala Met Gln Val Asp Tyr Ile Arg Val Tyr Lys Arg Asn	
150 355 360 365	
152 tga	1155

155 <210> SEQ ID NO: 3

156 <211> LENGTH: 384

157 <212> TYPE: PRT

158 <213> ORGANISM: Eisenia foetida

W--> 159 <400> SEQUENCE: 3

161 Met Arg Trp Thr Leu Val Val Cys Leu Leu Phe Gly Glu Gly Phe	
162 -15 -10 -5	
165 Ala Phe Thr Asp Trp Asp Gln Tyr His Ile Val Trp Gln Asp Glu Phe	
166 -1 1 5 10 15	
169 Asp Tyr Phe Asp Gly Ala Lys Trp Gln His Glu Val Thr Ala Thr Gly	
170 20 25 30	
173 Gly Gly Asn Ser Glu Phe Gln Leu Tyr Thr Gln Asp Gly Ala Asn Ser	
174 35 40 45	
177 Phe Val Arg Asp Gly Lys Leu Phe Ile Lys Pro Thr Leu Leu Ala Asp	
178 50 55 60	
181 Asn Ile Asn Pro Gln Thr Gly Ala Pro Phe Gly Thr Asp Phe Met Tyr	
182 65 70 75	
185 Asn Gly Val Leu Asp Val Trp Ala Met Tyr Gly Ala Cys Thr Asn Thr	
186 80 85 90 95	
189 Asp Asn Asn Gly Cys Tyr Arg Thr Gly Ala Ala Gly Asp Ile Pro Pro	
190 100 105 110	
193 Ala Met Ser Ala Arg Val Arg Thr Phe Gln Lys Tyr Ser Phe Thr His	
194 115 120 125	
197 Gly Arg Val Val Val His Ala Lys Met Pro Val Gly Asp Trp Leu Trp	
198 130 135 140	
201 Pro Ala Ile Trp Met Leu Pro Glu Asp Trp Val Tyr Gly Gly Trp Pro	
202 145 150 155	
205 Arg Ser Gly Glu Ile Asp Ile Ile Glu Thr Ile Gly Asn Arg Asp Phe	
206 160 165 170 175	
209 Lys Asn Thr Gly Glu Phe Leu Gly Ile Gln Lys Met Gly Ser Thr	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/596,101C

DATE: 02/12/2002
TIME: 11:43:16

Input Set: A:\EP.txt
Output Set: N:\CRF3\02122002\I596101C.raw

210	180	185	190
213	Met His Trp Gly Pro Gly Trp Asp Asp Asn Arg Tyr Trp Leu Thr Ser		
214	195	200	205
217	Leu Pro Lys His Ser Asp Asp Trp Asn Tyr Gly Asp Asn Phe His Thr		
218	210	215	220
221	Phe Trp Phe Asp Trp Ser Pro Asn Gly Leu Arg Phe Phe Val Asp Asp		
222	225	230	235
225	Glu Asn Gln Ala Leu Leu Asp Val Pro Tyr Pro Leu Ile Asp Ala Asn		
226	240	245	250
229	Pro Trp Trp Val Asp Phe Trp Glu Trp Gly Lys Pro Trp Leu Pro Gln		
230	260	265	270
233	Tyr Glu Asn Asp Asn Pro Trp Ala Gly Gly Thr Asn Leu Ala Pro Phe		
234	275	280	285
237	Asp Gln Asn Phe His Phe Ile Leu Asn Val Ala Val Gly Gly Thr Asn		
238	290	295	300
241	Gly Phe Ile Pro Asp Gly Cys Ile Asn Arg Gly Gly Asp Pro Ala Leu		
242	305	310	315
245	Gln Lys Pro Trp Ser Asn Gly Asp Trp Tyr Asn Asp Ala Met Arg Lys		
246	320	325	330
249	Phe Phe Asp Ala Arg Gly Asn Trp Lys Trp Thr Trp Asp Asp Glu Gly		
250	340	345	350
253	Asp Asn Asn Ala Met Gln Val Asp Tyr Ile Arg Val Tyr Lys Arg Asn		
254	355	360	365
257	<210> SEQ ID NO: 4		
258	<211> LENGTH: 33		
259	<212> TYPE: DNA		
260	<213> ORGANISM: Artificial Sequence		
262	<220> FEATURE:		
263	<223> OTHER INFORMATION: PCR Primer		
265	<220> FEATURE:		
266	<221> NAME/KEY: misc_feature		
267	<222> LOCATION: (2)..(2)		
268	<223> OTHER INFORMATION: Inosine		
270	<220> FEATURE:		
271	<221> NAME/KEY: misc_feature		
272	<222> LOCATION: (5)..(5)		
273	<223> OTHER INFORMATION: Inosine		
275	<220> FEATURE:		
276	<221> NAME/KEY: misc_feature		
277	<222> LOCATION: (8)..(8)		
278	<223> OTHER INFORMATION: Inosine		
280	<220> FEATURE:		
281	<221> NAME/KEY: misc_feature		
282	<222> LOCATION: (14)..(14)		
283	<223> OTHER INFORMATION: Inosine		
285	<220> FEATURE:		
286	<221> NAME/KEY: misc_feature		
287	<222> LOCATION: (25)..(25)		
288	<223> OTHER INFORMATION: Inosine		

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/596,101C

DATE: 02/12/2002
 TIME: 11:43:16

Input Set : A:\EP.txt
 Output Set: N:\CRF3\02122002\I596101C.raw

290 <220> FEATURE:
 291 <221> NAME/KEY: misc_feature
 292 <222> LOCATION: (28)..(28)
 293 <223> OTHER INFORMATION: Inosine
 295 <400> SEQUENCE: 4
 W--> 296 tnacngantg ggancaagta tcatngtntg gca 33
 299 <210> SEQ ID NO: 5
 300 <211> LENGTH: 32
 301 <212> TYPE: DNA
 302 <213> ORGANISM: Artificial sequence
 304 <220> FEATURE:
 305 <223> OTHER INFORMATION: PCR Primer
 307 <220> FEATURE:
 308 <221> NAME/KEY: misc_feature
 309 <222> LOCATION: (3)..(3)
 310 <223> OTHER INFORMATION: Inosine
 312 <220> FEATURE:
 313 <221> NAME/KEY: misc_feature
 314 <222> LOCATION: (6)..(6)
 315 <223> OTHER INFORMATION: Inosine
 317 <220> FEATURE:
 318 <221> NAME/KEY: misc_feature
 319 <222> LOCATION: (9)..(9)
 320 <223> OTHER INFORMATION: Inosine
 322 <220> FEATURE:
 323 <221> NAME/KEY: misc_feature
 324 <222> LOCATION: (12)..(12)
 325 <223> OTHER INFORMATION: Inosine
 327 <220> FEATURE:
 328 <221> NAME/KEY: misc_feature
 329 <222> LOCATION: (19)..(19)
 330 <223> OTHER INFORMATION: Inosine
 332 <400> SEQUENCE: 5
 W--> 333 aangtntgna anttgcacnc cgatagattc ca 32
 336 <210> SEQ ID NO: 6
 337 <211> LENGTH: 37
 338 <212> TYPE: DNA
 339 <213> ORGANISM: Artificial sequence
 W--> 340 <220> FEATURE:
 341 <223> OTHER INFORMATION: PCR Primer
 343 <400> SEQUENCE: 6
 344 ggggatccat gcattcacccg actgggatca atatcac 37
 347 <210> SEQ ID NO: 7
 348 <211> LENGTH: 29
 349 <212> TYPE: DNA
 350 <213> ORGANISM: Artificial sequence
 352 <220> FEATURE:
 353 <223> OTHER INFORMATION: PCR Primer
 355 <400> SEQUENCE: 7

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/596,101C

DATE: 02/12/2002

TIME: 11:43:17

Input Set : A:\EP.txt

Output Set: N:\CRF3\02122002\I596101C.raw

L:29 M:283 W: Missing Blank Line separator, <400> field identifier
L:159 M:283 W: Missing Blank Line separator, <400> field identifier
L:296 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:333 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:340 M:283 W: Missing Blank Line separator, <220> field identifier